

Multimedia for Art ReTrieval (M4ART)

Egon L. van den Broek^{a,b}, Thijs Kok^c, Theo E. Schouten^c, and Eduard Hoenkamp^b

^aDepartment of Artificial Intelligence, Vrije Universiteit Amsterdam,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

egon@few.vu.nl

<http://www.few.vu.nl/~egon/>

^bNijmegen Institute for Cognition and Information (NICI), Radboud University Nijmegen
P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

e.vandenbroek@nici.ru.nl

hoenkamp@acm.org

<http://eidetic.ai.ru.nl/egon/>

^cInstitute for Computing and Information Science (ICIS), Radboud University Nijmegen,
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

T.Kok@student.ru.nl

T.Schouten@cs.ru.nl

<http://eidetic.ai.ru.nl/thijs/>

<http://www.cs.ru.nl/~ths/>

ABSTRACT

The prototype for an online Multimedia for Art ReTrieval (M4ART) system is introduced, that provides entrance to the digitized collection of the National Gallery of the Netherlands (the Rijksmuseum). The current online system of the Rijksmuseum is text-based and requires expert knowledge concerning the work searched for, else it fails in retrieving it. M4ART extends this system by allowing the user to query with an example image that can be uploaded, or to select it by browsing the collection. The global color distribution of the example image, and perhaps a set of texture features, are then extracted and compared with those of the images in the collection. Thus, the collection can be queried based on text as well as content-based features. Moreover, the matching process of M4ART can be scrutinized. With this feature, M4ART not only integrates the means to give expert and lay equal access to the system, it also lets the user understand the system's inner workings. These qualities make M4ART unique in its ability to let the user access, enhance, and retrieve the knowledge available in digitized art collections.

Keywords: Content-Based Image Retrieval, Information Retrieval, art, color, texture, M4ART

1. INTRODUCTION

More and more digitized image archives are becoming available online (e.g., photo databases and museum collections). In the cultural domain, museums like the National Gallery of the Netherlands (the Rijksmuseum)¹ (see Figure 1), the Virtual Catalog for Art History,² and the Hermitage museum³ (see Figure 2) are extending their reach by making part of their collection available via Internet. To preserve our cultural heritage, many libraries, museums, and governmental institutes now follow these examples.

Making these collections publicly available calls for information systems to index, browse, and retrieve the information. However, most of these systems require expertise to use them effectively. This discourages laypeople to use the systems and explore the collections to their satisfaction.

This paper describes a Multimedia for Art ReTrieval (M4ART) system that gives entry to the collection of the Rijksmuseum. It provides suitable access to the database of art materials for both experts and laypersons. We start with a brief introduction that compares image retrieval by text and by content, and follow up explaining the domain of application in Section 2.1. The presentation of the results is an often ignored issue, which we will briefly touch in Section 3.1. In Section 3, the online M4ART system is introduced. We conclude this paper with a discussion of issues to be dealt with in future research.



Figure 1. A traditional text-based information retrieval system.

2. IMAGE RETRIEVAL: TEXT-BASED VERSUS CONTENT-BASED

In 1992, Kato⁴ introduced the term Content-Based Image Retrieval (CBIR), to describe his experiments on automatic retrieval of images from a database by color and shape features. Since then, CBIR has developed into a separate field.

CBIR is the application of computer vision to the image retrieval problem; i.e., the problem of searching for images in large image databases. Most image retrieval engines on the world wide web (WWW) make use of text-based image retrieval, in which images are retrieved based on their captions, descriptions, and surrounding text; e.g., see Figure 1. Although text-based image retrieval is fairly successful, it fully depends on the verbal annotations that accompany the images. Consequently, it requires every image in the database or on the WWW to be annotated.

A decade ago, Gudivada and Raghavan⁵ identified twelve fields of application in which CBIR can prove its usefulness: crime prevention, the military, intellectual property, architectural and engineering design, fashion and interior design, journalism and advertising, medical diagnosis, geographical information and remote sensing systems, cultural heritage, education, and training, home entertainment, and WWW searching. Despite this slew of applications, Smeulders, Worring, Santini, Gupta, and Jain⁶ noted in 2000 that “CBIR is at the end of its early years” and that it is certainly not the answer to all problems. In general, (i) CBIR techniques still yield unacceptable retrieval results, (ii) are restricted to the domain that is covered, (iii) lack a suitable user-interface, and (iv) are mainly technology-driven and, consequently, require the use of domain knowledge to fulfill the user’s information need.⁷

In the last decade,⁶⁻¹⁴ a change in research perspective with respect to CBIR systems can be seen: from computer vision and pattern recognition to other disciplines such as cognitive science and psychology. Hence, the issue of considering the human in the loop is gaining attention. Using knowledge about the user will provide insight in how the user-interface must be designed, how retrieval results may be presented, and what characterizes

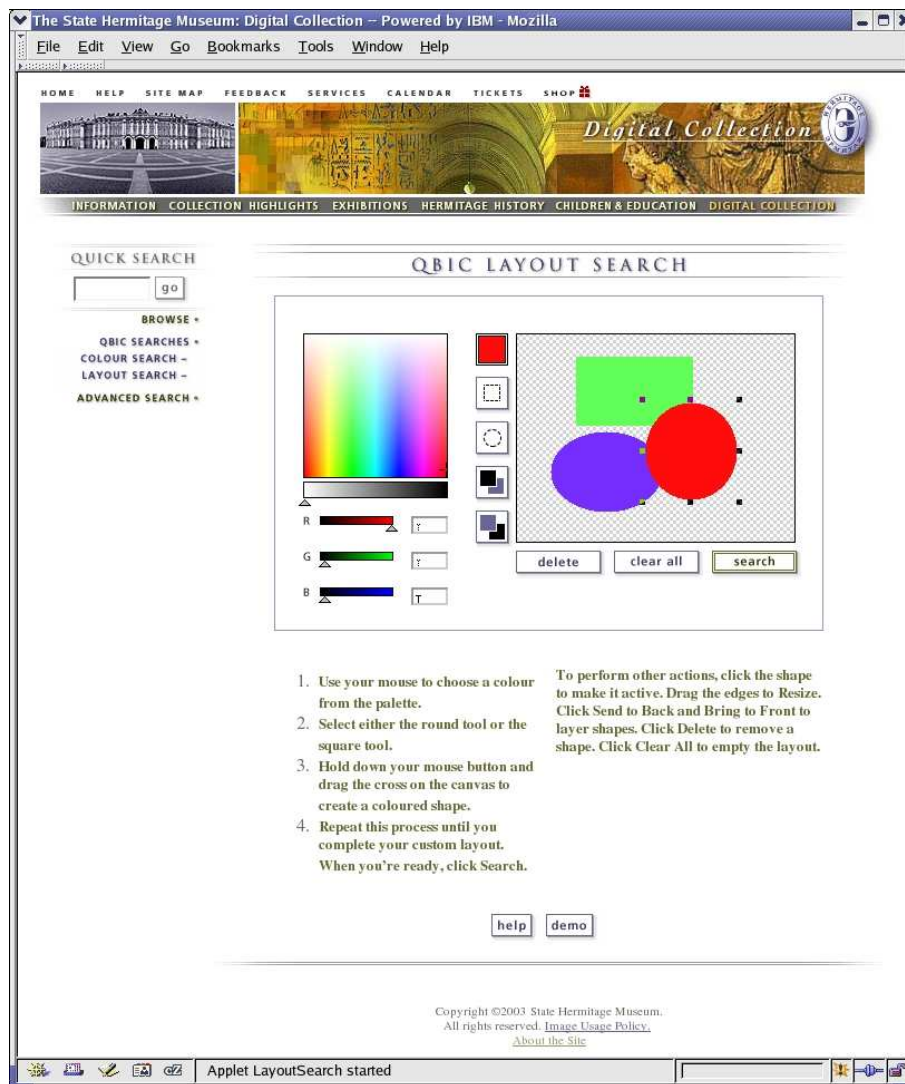


Figure 2. The interface of the Content-Based Image Retrieval as developed by IBM® for the Hermitage museum. It utilizes both color and shape in its processing scheme.

people's information need.¹⁵ Hence, in the line of research as discussed in this paper the human was constantly in the loop of technological development. The paradigm of “human-centered content-based image retrieval”, as introduced by Van den Broek,¹⁶ was adopted in the current research.

2.1. Domain of application

Historical archives are rich sources of information of which more and more are available in digitized form. For example, the institute for Dutch history (ING)¹⁷ recently finished a project with a time span of three decades. It recently introduced an online database of all correspondence of Willem of Orange. However, this project merely applies traditional information processing techniques. This paper uses the database of the National Gallery of the Netherlands (the Rijksmuseum). This database is already annotated. These annotations are already used in an online search system.¹

The Dutch Rijksmuseum states: “Research is the basic premise of all museum activity, from acquisition, conservation and restoration, to publication, education and presentation. In general, this research involves the object or work of art in the museum as being a source of information.”¹ However, how can these sources

of information be improved, and accessed effectively, and how can knowledge be retrieved from them? The Rijksmuseum has made their collection available online, through a web-interface,¹ as can be seen in Figure 1. Their current interface allows for traditional information retrieval; i.e., text-based search. Other recent initiatives are described in¹⁸⁻²⁰ amongst others. One of the best known CBIR systems available is IBM's QBIC,²¹ launched in the mid-1990s. In January 1997, the IBM Corporate Community Relations project with the State Hermitage Museum in St. Petersburg started.³ The goal of the project was to do much more than just provide technology to the Hermitage Museum. Its aim was to transform the very way people around the world experience the Hermitage Museum and its collections.³

With respect to the database of the Rijksmuseum, there is a more pressing problem than the proper use of keywords. In general, modern information retrieval techniques provide excellent results²² when two conditions are met: (i) a well annotated database is available and (ii) keywords are chosen that match the information need and also the keywords present in the database. For a modest sized database, such an approach can be highly successful. Not so for an ill-annotated database in an unrestricted domain, queried by non-professionals, using non-optimal keywords. And visitors of museum may not know the style of a painter, the period he lived in, or the title of the painting. A visitor may not even know the painter's name exactly. How to approximate a name, using keywords? In such a scenario, for the user there is little hope to fulfill his information-need.

A professional brings his knowledge to bear about the artist (such as name or country of origin), knowledge about the object (title, materials, techniques), dates, method of acquisition, and personal associations with the work of art. He can express detailed queries, which produce retrieval results with a high precision. However, how to define objects that evoke a similar atmosphere or trigger the same emotions? How to find objects with an alike expression, created using other techniques on different material? This calls for systems that are not confined to verbal expressions alone, and CBIR techniques may provide a much richer vocabulary to express an information need.

In the next section, we introduce the online Multimedia for Art ReTrieval (M4ART) system, which is available at: <http://cai.nici.ru.nl/M4ART/>. It integrates text and content-based image retrieval. Moreover, it provides a look under the hood, a view of its inner workings. Thus it meets two goals: (i) it provides access to the database for both laypeople and experts and (ii) it provides the means for the non-expert in the field of CBIR to gain an intuitive grasp of the CBIR techniques applied.

3. THE ONLINE MULTIMEDIA FOR ART RETRIEVAL (M4ART) SYSTEM

The online prototype Multimedia for Art ReTrieval (M4ART) system presented here consists of four components:

1. The two (inter)faces of M4ART
2. A feature extraction module
3. The matching engine
4. The SQL and image database, connected through the image IDs.

We will first discuss these components separately. Next, a new method is introduced that facilitates the user's understanding of how the system works.

3.1. The two (inter)faces of M4ART

In research on IR and CBIR systems, most effort is directed toward the underlying search technique. Often, the user interface of these systems receives little attention. In this respect, the current research differs from most other IR and CBIR research.

The M4ART incorporates two search engines: (i) A text-based search engine is included that fully relies on the annotations, as are present in the SQL database (see Figure 3) and (ii) A CBIR engine that utilizes features of the images present of the pieces of art; see Figure 4. Both engines provide a standard interface and an interface for advanced querying.

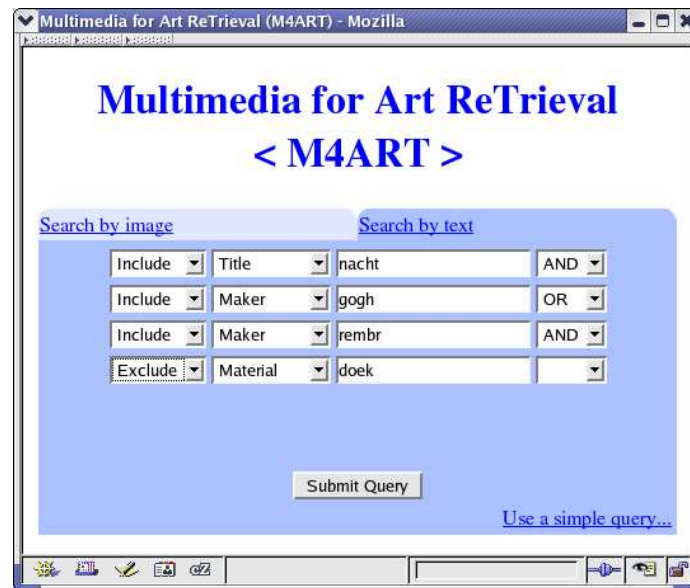


Figure 3. The advanced text-based image retrieval (Information Retrieval) interface of the online Multimedia for Art ReTrieval (M4ART) system. It provides the means for boolean queries. Each keyword has to be provided on a separate line. For each keyword the following parameters can be set: (i) whether the keyword should be included or excluded and (ii) which fields (i.e., ‘all fields’, title, artist, object, material, technique, and description) have to be queried for this keyword. In addition, (combinations of) conjunctions and disjunctions can be defined by the connectives ‘and’ and ‘or’. The M4ART system is available at <http://cai.nici.ru.nl/M4ART/>.

3.1.1. Text-based art retrieval

The standard interface provides one line to access keywords. All database fields (i.e., ‘all fields’, title, artist, object, material, technique, and description) of all images in the database are queried for the keywords provided. The advanced interface provides the means for boolean queries. Each keyword has to be provided on a separate line. For each keyword the following parameters can be set: (i) whether the keyword should be included or excluded and (ii) which fields have to be queried for this keyword. In addition, (combinations of) conjunctions and disjunctions can be defined by the connectives ‘and’ and ‘or’; see Figure 3.

After a text or content-based query, the results can be inspected, using content-based techniques. In a matrix, the distances between all possible pairs of retrieved images are provided, as is shown in Figure 5b. Moreover, when the mouse pointer is placed upon a distance, the two images to which the distance refers to are shown. Hence, the relation between all images retrieved can be inspected. Consequently, the users can infer the system’s working. Subsequently, users can learn to understand and even respect the system’s performance.

3.1.2. Content-Based Art Retrieval (C-BAR)

At the XVIth International Conference of the Association of History and Computing, the notion Content-Based Art Retrieval (C-BAR) was introduced.²³ C-BAR is simply CBIR applied on the domain of art, where characteristics of a piece of art or a digital photo of an art object defines the query. In M4ART, an image can be provided in three ways to the engine: (i) the catalog can be browsed for an image, (ii) the URL of an image can be provided, and (iii) an image can be uploaded from the computer of the user; see Figure 4.

De Greef and Van Eijk²⁴ determined that users prefer to browse image databases manually. Moreover, they determined that a grid of 9–16 images per screen is optimal for visualizing an image collection. Another interesting finding was that the difference in size between the (real-size) image (e.g., a painting) that was seen and its thumbnail version, as it is available in the image database, does not hinder recognition. In order to facilitate efficient and user friendly browsing, the user interfaces are designed conform the guidelines of these authors²⁴ and provides either 12 or 15 images per screen as is shown in Figure 5a and Figure 3.2.2.



Figure 4. The advanced Content-Based Image Retrieval (CBIR) interface of the online Multimedia for Art ReTrieval (M4ART) system. An image can be selected through browsing the database, providing the URL of an image, or by uploading an image from the local hard-disk. A color space and its quantization scheme can be chosen; i.e., HSI $18 \times 3 \times 3$, 11 color categories, RGB $4 \times 4 \times 4$, or HSI $9 \times 6 \times 6$ (see Section 3.2.1). A choice can be made between the intersection and Euclidean distance measure (see Section 3.3). Texture features can be selected to be taken in the processing scheme (see Section 3.2.2). The M4ART system is available at <http://cai.nici.ru.nl/M4ART/>.

A image conversion module is in M4ART included. This module can convert most image formats (e.g., gif, jpeg, jpeg-2000, png, bmp, tiff) and some other formats (e.g., html, eps, ps, and pdf) to jpeg, which is the image format used for the database. The database is incrementally updated with the new images provided by the users.

3.2. Feature extraction module

With features we denote characteristics of objects. In the context of images, features can denote colors, texture descriptors, and shape descriptors. In the current research, we use features derived from the global color distribution of the image optionally combined with features that describe the texture properties of the image.

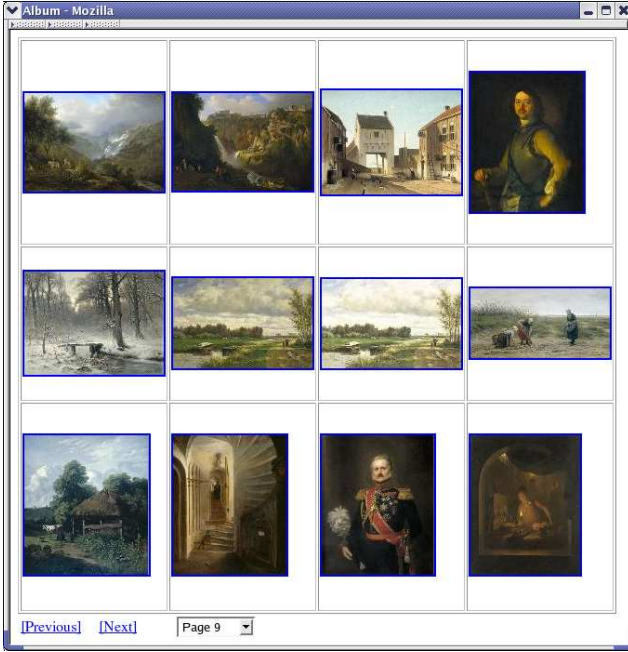
3.2.1. Color

A color space specifies colors as tuples of (typically three) numbers, conform to certain specifications. For image processing purposes, color spaces are often quantized. The color space in which this is done determines the perceptual intuitivity of the quantization up to a high extend. Moreover, the axes of the color space can be quantized, using a different scheme for each axis. Again, this depends on the color space of choice.

A color space is perceptually intuitive if distances between points in that space (i.e., ‘colors’) have a relation to perceived closeness of these ‘colors’ by human observers. If that relation is constant one can even speak of perceptual uniformity. In this section, we describe the color spaces used and the quantization schemes applied on them. The quantization of color images transformed into gray-scale images will not be described for every color space since it is the same for every color space: the gray-scale axis is divided in the number of bins needed for the specific quantization scheme.

A quantization scheme provides the means to determine a color histogram. Such a histogram can be determined for parts of the image as well as for the image as a whole. The latter application of the color histogram is applied in the current research. It describes the global color characteristics of an image.

The RGB (Red, Green, and Blue) color space is the most used color space for computer graphics and is not perceptually uniform. Each color-axis (R, G, and B) is equally important and is quantized with the same



(a)



(b)

Figure 5. The online Multimedia for Art ReTrieval (M4ART) system, available at <http://cai.nici.ru.nl/M4ART/>. (a) Browsing through the database of the Rijksmuseum. On each page 12 pieces of art are shown. Both browsing page by page as selecting a page directly are provided. (b) The comparison of text-based and content-based results. Of all retrieved images by the text-based query the content-based distances are provided in a matrix. When the user hovers over the distances, the images between which the distance is determined are shown.

precision ($4 \times 4 \times 4$). The conversion from a RGB image to a gray value image simply takes the sum of the R, G, and B values and divides the result by three.

The HSI (Hue, Saturation, and Value) color space is more closely related to human color perception than the RGB color space²⁵ and is perceptually intuitive but not perceptually uniform. Hue is the color component of the HSI color space. When Saturation is set to 0, Hue is undefined and the Value-axis represents the gray-scale image. Two quantizations of HSI are used: 162 ($18 \times 3 \times 3$) and 324 ($9 \times 6 \times 6$) bins.

Another view on color representation is the concept of 11 color categories (i.e., black, white, red, green, yellow, blue, brown, purple, pink, orange, and gray), as introduced by Berlin and Kay.²⁶ Since then, several researchers discussed the topic; see Derefeldt et al.²⁷ for an overview. Van den Broek et al.²⁸ developed a method to describe the complete HSI color space, based on a limited set of experimentally determined, categorized colors. This method provided a unique color space segmentation, which can be applied as an 11 color categories, quantization scheme.

3.2.2. Texture

Before texture can be analyzed, a color quantization scheme has to be applied, as discussed in the previous section. Next, several texture analysis techniques can be applied, both for general and for specific purposes. We have chosen for one of the more intuitive texture descriptors: the color correlogram,^{29, 30} a sequential color-based texture analysis method: first color is quantized and second texture is analyzed.

The color correlogram $C_{\bar{d}}(i, j)$ counts the co-occurrence of pixels with color values i and j at a given distance \bar{d} . The distance \bar{d} is defined in polar coordinates (d, α) , with discrete length and orientation. In practice, α takes

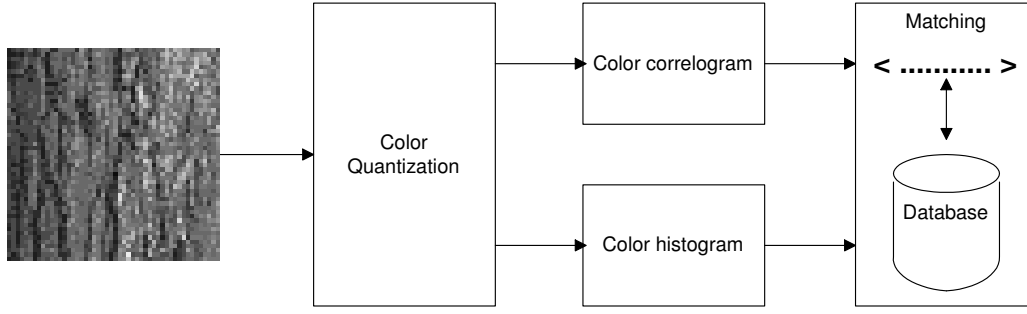


Figure 6. The parallel-sequential approach for texture analysis which yields in parallel: global color analysis, using the color histogram, and color induced texture analysis, using the color correlogram.

the values $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ$, and 315° . The color correlogram $C_{\bar{d}}(i, j)$ can now be defined as:

$$C_{\bar{d}}(i, j) = \Pr(I(p_1) = i \wedge I(p_2) = j \mid |p_1 - p_2| = \bar{d}), \quad (1)$$

where \Pr is probability, and p_1 and p_2 are positions in the gray-scale image I .

The algorithm yields a symmetric matrix; hence, only angles up to 180° need to be considered. A single co-occurrence matrix can be defined for each distance d by averaging four co-occurrence matrices of different angles (i.e., $0^\circ, 45^\circ, 90^\circ$, and 135°).

Because of the high dimensionality of the matrix, the individual elements of the color correlogram are rarely used directly for texture analysis. Instead, textural features can be derived from the matrix. In previous research,³¹ an optimal feature-distance combination was determined, derived from the color correlogram. The best classification was found using a combination of four features: entropy, inverse difference moment, cluster prominence, and Haralick's correlation, with $d = 1$. Consequently, this configuration was chosen for this research.

In the user-interface for advanced query schemes, the user can choose to include texture features in the matching process. Recently, this combination of the color histogram and the color correlogram proved to provide high classification results.³¹ This way of analyzing color images is named: parallel-sequential texture analysis. Figure 3.2.2 provides the processing scheme of parallel-sequential texture analysis.

3.3. The matching engine

The matching engine has an interface that connects to the query definition user interface and to the user interface that presents the retrieved (matching) results. Through the query definition user interface, the matching engine receives the query image as well as the necessary parameters: the feature vector to be used and the number of results to be presented. Subsequently, the matching engine extracts the features from the query image and matches it to the image features extracted from the images in the database. The results that match are ranked based on their distance to the query image. Next, the images that match best are sent to the user interface that presents the results, as is presented in Figure 3.2.2.

The query image has to be compared with the images in the database; i.e., the feature vectors of the images are compared with each other. In order to express the (dis)similarity of two features vectors, two distance metric can be used. In literature, a wide variety of distance measures can be found. Both dissimilarity measures for feature vectors are based upon the Minkowski metric: the intersection distance and the Euclidean distance. The Minkowski metric between two vectors v and w is defined as:

$$d^k(v, w) = \left(\sum_{i=0}^{I-1} |v[i] - w[i]|^k \right)^{\frac{1}{k}}, \quad (2)$$

where I denotes the number of elements of the vector. For the intersection distance $k = 1$ and for the Euclidean distance $k = 2$.



Figure 7. The result of M4ART after a standard CBIR query. The query image was selected from the database of the Rijksmuseum.

Note that the distance measures used, require the normalization of vectors. This can be achieved either by normalizing all images to an arbitrary size before they are processed or by normalizing vector v that resulted from the processed image. For the latter, we introduce the following class of normalizations:

$$v^k = \frac{v}{(\sum_{i=0}^{I-1} |v[i]|^k)^{1/k}}. \quad (3)$$

3.4. Technical specifications

So far this section described the three components that can be identified in M4ART. In this part, we will present the implementation that is build based on this design. M4ART relies on a popular combination of open source programming facilities; i.e., Apache, PHP and MySQL. All applications run on a Redhat Linux server. These programs are beneficial in several ways: the Apache webserver guarantees the great amount of traffic is handled correctly, PHP provides numerous functions that allow effective development and MySQL offers a way to store and retrieve large amounts of data.

The Rijksmuseum uses an Adlib[®] information systems for archives, museums and libraries in which the paintings were annotated. The annotations were connected to the images themselves through the ID of the images. The Rijksmuseum provided us with the database in XML format; however, the conversion from Adlib to XML format resulted in numerous errors. After these errors were semi-automatically removed, the XML database was used in the initial version of M4ART.²³

For the current version of M4ART, we wanted to include the means to do Boolean searches and wanted to increase the speed of searching. The use of an SQL database provided the means to fulfill both aims: Boolean queries became possible and the speed of searching increased with the use of the SQL database. The conversion from XML to the SQL database did not cause any problems. Additionally, the museum provided us with the image database itself, containing an approximate of 3400 JPG images. The color- and texture information had to be derived from these images in order to provide the CBIR functionality. A Java utility was designed to process all the images and save the relevant data into a file. Due to the fact that different color models could be used to derive the information, multiple files were created. The amount of data gathered during this parsing sessions was huge: every file that was created contained a record of every image including its characterization, resulting in file sizes exceeding 30 MB.

Since the SQL implementation proved to be a success, the characterization of the images - the histograms and the texture features - were also converted to the SQL database. At that point, the database served as a resource facility, in which annotation as well as characterization could be stored as well as retrieved. An extra advantage was gained due to the fact that CBIR and IR information could be combined: for example, the results of a Boolean query can be compared to CBIR.

The MySQL interface is integrated in the PHP extension; this enabled the implementation of queries and parsing results. In the initial design, the file reading and parsing procedures were obsolete; in the current system, the SQL statements are used. The evaluation of the distance between the vectors is also implemented using PHP scripts. Existing functions for mathematical operations as well as the MySQL interface allowed a clear structure of the source code, rapid development and overall fast performance. PHP is also responsible for building the user interface of M4ART. The HTML code that is produced complies with the HTML 4.0x specification of the WWW consortium (W3C), which guarantees that the most common browsers display the pages in the same manner. The layout of the interface is clean and straightforward. The website can be viewed using different screen resolutions since the interface elements are automatically resized based on the resolution of the user's screen.

With the latter specifications, all aspects of the M4ART system are discussed. In the next section, we will end this paper with some conclusions and plans for future work.

4. CONCLUSION

With M4ART an online system is launched that provides access to the art collection of the Rijksmuseum. Both text-based and content-based queries can be done. Both can be utilized in their standard form and can be specified in an advanced interface. M4ART enables access for both experts and laypersons. Moreover, it provides insight information concerning the process of content-based access.

In the continuation of the current research, all aspects of the system will be improved. The user interface will be enhanced both with respect to its visual appearance and with respect to its functionality. In general, for complex (multi-media) data-mining systems, user profiling is adopted as a paradigm. In contrast, we propose to include artist profiling instead. Subsequently, features as can be found in artists' work are identified (e.g.,

number of colors, contrasts, texture) and utilized in such a profile. The resulting prototype feature vector could enhance the M4ART system substantially.

In time, ontology-based knowledge representations should be used to define structure in the database instead of merely flat SQL data representation. Moreover, cooperative annotation as envisioned and applied by Schomaker, Vuurpijl, and De Leau³² should be applied. Of special interest would be to compare how laypersons and experts annotate the material. Learning algorithms (e.g., neural networks) or techniques as singular value decomposition can extract underlying (implicit) dimensions of judgment,²² which are common to both groups. Moreover, specific features can be provided based on the domain knowledge of the user.

M4ART is still in development; however, its current version showed to work successful. It provides access both to an advanced text-based search engine for experts and to (ten configurations for) content-based retrieval that do not require any domain knowledge. Moreover, the means to understand its working are provided. Hence, a unique system is introduced to access, enhance, and retrieve knowledge from digitized art collections.

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO) under project numbers 634.000.001 and 634.000.018. We thank Xenia Henny and Kees Schoemaker for their cooperation. They provided us the database of the Rijksmuseum.

REFERENCES

1. Rijksmuseum, "Research," URL: <http://www.rijksmuseum.nl/wetenschap/> [Last accessed on October 24, 2005].
2. U. L. Karlsruhe, "Virtual catalogue for art history (vkk)," URL: http://www.ubka.uni-karlsruhe.de/kvk/vkk/vk_kunst_engl.html [Last accessed on October 24, 2005].
3. IBM, "The state hermitage museum," URL: <http://www.heritagemuseum.org/> [Last accessed on October 24, 2005].
4. T. Kato, "Database architecture for content-based image retrieval," in *Proceedings of SPIE Image Storage and Retrieval Systems*, A. A. Jambardino and W. R. Niblack, eds., **1662**, pp. 112–123, (San Jose, CA, USA), February 1992.
5. V. N. Gudivada and V. V. Raghavan, "Content-based image retrieval systems," *IEEE Computer* **28**(9), pp. 18–22, 1995.
6. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(12), pp. 1349–1380, 2000.
7. Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Past, present, and future," *Journal of Visual Communication and Image Representation* **10**, pp. 1–23, 1999.
8. L. H. Armitage and P. G. B. Enser, "Analysis of user need in image archives," *Journal of Information Science* **23**(4), pp. 287–299, 1997.
9. G. Ciocca and R. Schettini, "Using a relevance feedback mechanism to improve content-based image retrieval," in *Visual Information and Information Systems: Third International Conference (VISUAL'99)*, D. P. Huijsmans and A. W. M. Smeulders, eds., *Lecture Notes in Computer Science* **1614**, pp. 107–114, Springer-Verlag GmbH, 1999.
10. N. Ikonomakis, K. N. Plataniotis, and A. N. Venetsanopoulos, "User interaction in region-based color image segmentation," in *Visual Information and Information Systems: Third International Conference (VISUAL'99)*, D. P. Huijsmans and A. W. M. Smeulders, eds., *Lecture Notes in Computer Science* **1614**, pp. 99–106, Springer-Verlag GmbH, 1999.
11. M. Israël, E. L. van den Broek, P. van der Putten, and M. J. den Uyl Visual Alphabets: Video Classification by End Users, in *Multimedia Data mining and Knowledge Discovery*, ch. 12 (Part III), V. A. Petrushin and L. Khan, eds., Springer-Verlag: Berlin - Heidelberg, [in press].
12. C. Jørgensen, "Access to pictorial material: A review of current research and future prospects," *Computers and the Humanities* **33**(4), pp. 293–318., 1999.

13. T. P. Minka, "An image database browser that learns from user interaction," Tech. Rep. 365, MIT Media Laboratory, 1996.
14. L. Vuurpijl, L. Schomaker, and E. L. van den Broek, "Vind(x): Using the user through cooperative annotation," in *Proceedings of the Eighth IEEE International Workshop on Frontiers in Handwriting Recognition*, S. N. Srihari and M. Cheriet, eds., pp. 221–226, IEEE Computer Society, Los Alamitos, CA, (Ontario, Canada), 2002.
15. E. C. M. Hoenkamp, O. Stegeman, and L. R. B. Schomaker, "Supporting content retrieval from www via "basic level categories"," in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 311–312, ACM Press: New York, NY, USA, 1999.
16. E. L. van den Broek, *Human-Centered Content-Based Image Retrieval*. PhD thesis, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, 2005; Online available: <http://eidetic.ai.ru.nl/egon/PhD-Thesis/> [Last accessed on October 24, 2005].
17. Instituut voor Nederlandse Geschiedenis, "Website Instituut voor Nederlandse Geschiedenis (ING)," URL: <http://www.inghist.nl> [Last accessed on October 24, 2005].
18. F. Hernandez, C. Wert, I. Recio, B. Aguilera, W. Koch, M. Bogensperger, P. Linde, G. Günter, B. Mulrenin, X. Agenjo, R. Yeats, L. Bordoni, and F. Poggi, "XML for libraries, archives, and museums: The projects COVAX," *Applied Artificial Intelligence* **17**(8), pp. 797–816, 2003.
19. N. Davenport, "Council on library and information resources." URL: <http://www.clir.org/> [Last accessed on October 24, 2004].
20. J. Trant, *Image Retrieval Benchmark Database Service: A Needs Assessment and Preliminary Development Plan*, Archives & Museum Informatics, Canada, 2004.
21. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC system," *IEEE Computer* **28**(9), pp. 23–32, 1995.
22. E. C. M. Hoenkamp, "Unitary operators on the document space," *Journal of the American Society for Information Science and Technology* **54**(4), pp. 319–325, 2003.
23. E. L. van den Broek, T. Kok, E. Hoenkamp, Th. E. Schouten, P. J. Petiet, and L. G. Vuurpijl, Content-Based Art Retrieval (C-BAR), in *Humanities, Computers and Cultural Heritage*, O. Boonstra, L. Beurne, P. Doorn, J. van den Herik, B. de Nil, and P. Witkamp, eds., pp. 70–77. Amsterdam – The Netherlands: Royal Academy of Arts and Sciences.
24. H. P. de Greef and R. L. J. van Eijk, "Visually searching a large image database: Manual browsing versus rapid visual presentation," [In preparation].
25. T. Lin and H. Zhang, "Automatic video scene extraction by shot grouping," in *Proceedings of the 15th IEEE International Conference on Pattern Recognition*, **4**, pp. 39–42, (Barcelona, Spain), 2000.
26. B. Berlin and P. Kay, *Basic color terms: Their universals and evolution*, Berkeley: University of California Press, 1969.
27. G. Derefeldt, T. Swartling, U. Berggrund, and P. Bodrogi, "Cognitive color," *Color Research & Application* **29**(1), pp. 7–19, 2004.
28. E. L. van den Broek, P. M. F. Kisters, and L. G. Vuurpijl, "Content-based image retrieval benchmarking: Utilizing color categories and color distributions," *Journal of Imaging Science and Technology* **49**(3), pp. 293–301, 2005.
29. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *Transactions on Systems, Man and Cybernetics* **3**(6), pp. 610–621, 1973.
30. J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, G. Medioni, R. Nevatia, D. Huttenlocher, and J. Ponce, eds., pp. 762–768, 1997.
31. E. L. van den Broek and E. M. Rikxoort, "Parallel-sequential texture analysis," *Lecture Notes in Computer Science (Advances in Pattern Recognition)* **3687**, pp. 532–541, 2005.
32. L. Schomaker, L. Vuurpijl, and E. de Leau, "New use for the pen: outline-based image queries," in *Proceedings of the 5th IEEE International Conference on Document Analysis*, pp. 293–296, (Piscataway (NJ), USA), 1999.